



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ
ГОСУДАРСТВЕННАЯ РЕГИСТРАЦИЯ ПРОГРАММЫ ДЛЯ ЭВМ

Номер регистрации (свидетельства):
2020667189

Дата регистрации: 21.12.2020

Номер и дата поступления заявки:
2020666598 14.12.2020

Дата публикации и номер бюллетеня:
21.12.2020 Бюл. № 1

Контактные реквизиты:

Телефон: +7 (499) 263-69-84, E-mail:
patent.bmstu@mail.ru

Автор(ы):

Бородулин Алексей Сергеевич (RU),
Нелюб Владимир Александрович (RU),
Гантимуров Андрей Павлович (RU),
Соколов Григорий Владимирович (RU)

Правообладатель(и):

Федеральное государственное бюджетное
образовательное учреждение высшего
образования "Московский государственный
технический университет имени Н.Э. Баумана
(Национальный исследовательский
университет)" (RU)

Название программы для ЭВМ:

Модуль препроцессинга неструктурированного текста с применением технологий искусственного интеллекта

Реферат:

Программа предназначена для очистки текста от неинформативных признаков, таких как знаки препинания, предлоги, специальные символы и подобные элементы, в которых отсутствует информативность. Также программа осуществляет приведение к нижнему регистру шрифта, лемматизацию текста и токенизацию, создание словаря на основе модели "Bag of Words". В основе программы лежат алгоритмы искусственного интеллекта. Тип ЭВМ: IBM PC-совмест. ПК. ОС: Linux.

Язык программирования: Python

Объем программы для ЭВМ: 597 КБ